Comment

emotion review

Emotion Review
Vol. 2, No. 3 (July 2010) 288–289

ISSN 1754-0739
DOI: 10.1177/1754073910368735
er.sagepub.com

Latent and Emergent Models in Affective Computing

Rafael A. Calvo
School of Electrical and Information Engineering, The University of Sydney, Australia

Abstract

New research on affective computing aiming to develop computer systems that recognize and respond to affective states can also contribute to the issues raised by Coan. Research on how humans interact with computers, and computer models that automatically recognize affective states from features in our physiology, behaviour, and language, may provide insights on how emotions that are experienced and expressed come to be. For example, there is empirical evidence that affect recognition techniques using several modalities are more accurate than those using a single modality, but it is not clear if these improvements are caused by superadditivity (i.e., emergence) or redundancy.

Keywords
affective computing, affective states, computer models

In Coan's (2010) discussion of latent and emergent models he notes the recent growth of the area of emotion research, with people in disciplines as varied as sociology and economics taking part in this study of the human experience. However, at least one group of researchers remains conspicuously absent from his list, and I would suggest this group is well worth this humble commentary, as it could also prove to be a singularly invaluable ally to the ongoing work of psychologists and other researchers in the area of emotion research. Computer scientists and engineers have notably turned their attention to the study of emotions and affective science, particularly in the last 10 years. The very models Coan discusses are exemplified in the work and assumptions made by these researchers.

''Affective computing'' (Picard, 1997), generally looks into ways to make technological artifacts more emotionally intelligent, that is, be able to recognize (e.g., from a person's facial expressions), respond to (e.g., adapting the interface) and represent (e.g., in avatars) affective states. The purpose is to make the use of computer technologies more productive and enjoyable.

Most affective computing research has been based on the information processing approach, where stimuli (generally produced by the interaction with a computer) affect a subject, triggering an emotional state that, in principle, can be measured without loss of information through one or more sensing modalities (speech, facial expression, physiology, etc.). This is what Coan refers to as the latent variables model.

Despite being at the core of most human–computer interaction design, the information processing approach has been questioned in recent years (Dourish, 2004), particularly in the area of affective interfaces (Boehner, DePaula, Dourish, & Sengers, 2005). Most of these critics argue that the information processing approach, and its reductionist approaches, cannot explain the complexity of embodied and social experiences.

A general assumption in affective computing has been that measuring more modalities (facial expressions, movement, speech, etc.) will increase the accuracy of recognition systems (Pantic, Sebe, Cohn, & Huang, 2005). The latent model indicates that measuring many modalities should not be necessary, as they are inherently coherent, yet when different modalities are combined evidence suggests that classification accuracy increases (Pantic et al., 2005).

The improvements in classification accuracy may be due to redundancy in the data or from superadditive properties (D'Mello & Graesser, 2007). One way of measuring the emergent properties of combining multiple modalities is to measure the level of superadditivity. In order for the accuracy to be superadditive the kappa statistic of the combined model should be statistically higher than the individual ones.

Empirical evidence on how data fusion improves accuracy is not clear. D'Mello and Graesser (2007, in press) recorded subjects while they interacted with a computer tutor. The interactions were then annotated with information on the affective states of boredom, confusion, flow (engagement), and frustration, reported by different judges (learners themselves, untrained peers, and trained judges). Using one metric of superadditivity used in some neuroscience studies (Meredith & Stein, 1983), the study showed a low level (0.3) of superadditivity when combining information from these two sensors and labels produced by self-reports and reports by peers (but not so when data was

Corresponding author: Rafael A. Calvo, School of Electrical and Information Engineering, Building J03, Maze Crescent, The University of Sydney, NSW 2006, Australia. *Email*: rafael.calvo@sydney.edu.au

annotated by trained judges). Using another measure of superadditivity (King & Palmer, 1985), the study points to improvements being due to redundancy instead of superadditivity.

D'Mello and Graesser's (2007) study is prototypical of affective computing research. An underlying methodological challenge in these studies is that they normally rely in self- or third-party reports (or a combination of both) to create the training set used to build the affect discrimination models. This is highlighted above, by the dependence of the results on who annotated the recording (self, peer, or expert).

In another type of affective computing applications the aim is to build computer systems that express emotions. Significant work in this area is based on the idea that avatars or robots that express emotions can communicate more efficiently with users, for example in the area of assisted care. In one such study, Cañamero (2005) studied robots that followed simple behavioral rules. Despite the simplicity of the rules, when the robots moved, humans labeled the robots' behavior as ''emotional,'' and the authors described these emotions as emergent phenomena. This work shows a way in which a combination of simple behavioral rules, and the human tendency to anthropomorphize, can lead to emergent phenomena that humans label as emotional. It also displays again the issue of how subjective the labeling of emotions can be.

References

Boehner, K., DePaula, R., Dourish, P., & Sengers, P. (2005). Affect: From information to interaction. In O. W. Bertelsen et al. (Eds.), *4th Decennial Conference on Critical Computing: Between Sense and Sensibility* (pp. 59–68). Aarhus, Denmark: ACM.

Cañamero, L. (2005). Emotion understanding from the perspective of autonomous robots research. *Neural Networks*, *18*, 445–455.

Coan, J. (2010). Emergent ghosts of the machine. *Emotion Review*, *2*(3), 274–285.

D'Mello, S., & Graesser, A. (2007). Mind and body: Dialogue and posture for affect detection in learning environments. In R. Luckin et al. (Eds.), *Artificial intelligence in education* (pp. 161–168). Los Angeles, CA: IOS Press.

D'Mello, S., & Graesser, A. (in press). Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-adapted Interaction*.

Dourish, P. (2004). *Where the action is: The foundations of embodied interaction*. Cambridge, MA: MIT Press.

King, A. J., & Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research*, *60*, 492–500.

Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science*, *221*(4608), 389–391.

Pantic, M., Sebe, N., Cohn, J. F., & Huang, T. (2005). Affective multimodal human-computer interaction. In T. Chua & H. Zhang (Eds.), *13th Annual ACM International Conference on Multimedia* (pp. 669–676). Singapore: ACM.

Picard, R. (1997). *Affective computing*. Cambridge, MA: MIT Press.